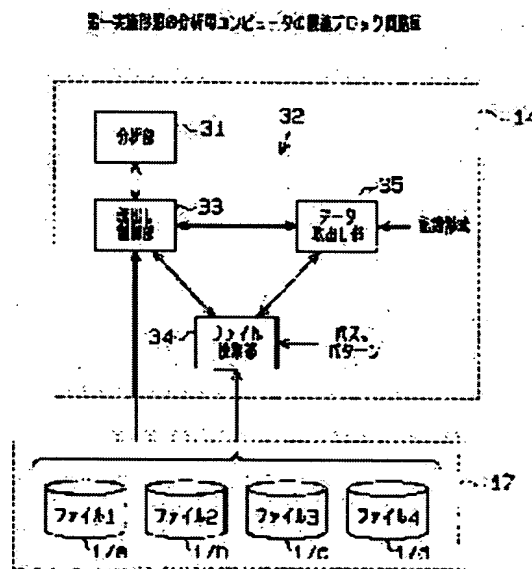


(11)Publication number : 2001-142748  
(43)Date of publication of application : 25.05.2001

(71)Applicant : FUJITSU LTD  
(72)Inventor : MORITA KATSUYA  
ITO MASARU  
YAMAZAKI YOSHIHIRO

**SOLUTION:** A data readout part 32 is equipped with a data extraction part 35 which extracts desirable data including time information from log data of respective files 17a to 17d, a file retrieval part 34 which retrieves a file to be processed next by sorting pieces of time information that the extracted head log data of the files 17a to 17d by the data extraction part 35 has in time series, and a readout control part 33 which controls the file retrieval part 34 and data extraction part 35 so that the log data are read out in time series.



[Date of request for examination]  
[Date of sending the examiner's decision of rejection]  
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]  
[Date of final disposal for application]  
[Patent number]  
[Date of registration]  
[Number of appeal against examiner's decision of rejection]

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2001-142748

(P2001-142748A)

(43)公開日 平成13年 5月25日 (2001.5.25)

(51)Int.Cl. <sup>7</sup>	識別記号	F I	テマコード <sup>*</sup> (参考)
G 0 6 F 11/34		G 0 6 F 11/34	S 5 B 0 4 2
12/00	5 2 0	12/00	5 2 0 E 5 B 0 7 5
	5 4 6		5 4 6 P 5 B 0 8 2
17/30		15/40	3 7 0 Z

審査請求 未請求 請求項の数 8 O L (全 9 頁)

(21)出願番号 特願平11-326878

(22)出願日 平成11年11月17日(1999.11.17)

(71)出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番1号

(72)発明者 森田 賀津也

愛知県名古屋市中区葵1丁目16番38号 株式会社富士通愛知エンジニアリング内

(72)発明者 伊藤 勝

愛知県名古屋市中区葵1丁目16番38号 株式会社富士通愛知エンジニアリング内

(74)代理人 100068755

弁理士 恩田 博宣 (外1名)

最終頁に続く

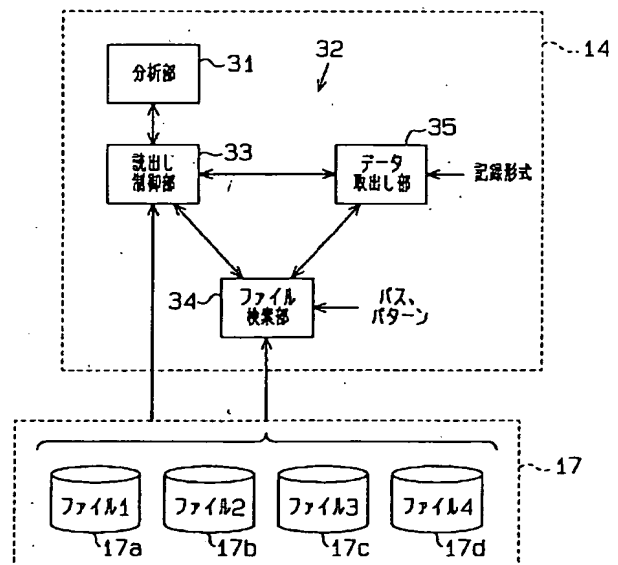
(54)【発明の名称】 ログ読出し方法、ログ読出し装置、分析装置、及び記録媒体

(57)【要約】

【課題】格納媒体が複数ファイルで構成されるログについて、サーバ固有の格納媒体生成規則に依存しない、汎用的なログのログ読出し装置を提供すること。

【解決手段】データ読出し部32は、各ファイル17a～17dのログデータから時刻情報を含む所望のデータを取り出すデータ取出し部35と、データ取出し部35により取り出された各ファイル17a～17dの先頭ログデータが持つ時刻情報を時系列にソートして次に処理するファイルを検索するファイル検索部34と、ファイル検索部34及びデータ取出し部35を、ログデータを時系列に沿って読み出すべく制御する読出し制御部33とを備えた。

第一実施形態の分析用コンピュータの機能ブロック図



BEST AVAILABLE COPY

## 【 特許請求の範囲】

【 請求項1 】 複数のファイルに格納されたログデータを時系列に沿って読み出すログ読出し方法において、前記各ファイルの先頭ログデータが持つ時刻情報をそれぞれ取出し、該時刻情報を時系列でソートして前記複数のファイルの読出し順序を決定する、ことを特徴とするログ読出し方法。

【 請求項2 】 複数のファイルに格納されたログデータを時系列に沿って読み出すログ読出し装置において、前記各ファイルのログデータから時刻情報を含む所望のデータを取出すデータ取出し部と、前記データ取出し部により取り出された各ファイルの先頭ログデータが持つ時刻情報を時系列にソートして次に処理するファイルを検索するファイル検索部と、前記ファイル検索部及びデータ取出し部を、前記ログデータを時系列に沿って読み出すべく制御する読出し制御部と、を備えた、ことを特徴とするログ読出し装置。

【 請求項3 】 前記読出し制御部は、1つのファイルに対する読み出しが終了し、次のファイルを決める段階でその都度前記ファイル検索部を動作させ、前記複数のファイルをソートして読出し順序を決定する、ことを特徴とする請求項2に記載のログ読出し装置。

【 請求項4 】 前記ファイル検索部は、前記複数のファイルのうちの2つを順次オープンし、双方のアクセス環境を保持した状態でそれぞれのファイルの先頭ログが持つ時刻情報を比較し、その比較結果に基づいて前記読出し順序を決定する、ことを特徴とする請求項2に記載のログ読出し装置。

【 請求項5 】 前記読出しの中断時またはログデータの終了時の情報を履歴ファイルに保存する履歴記録部を備え、前記読出し制御部は、前記履歴ファイルから読み出した情報に基づいて前記中断時または終了時の読出し位置を復元して読出しを再開する、ことを特徴とする請求項2に記載のログ読出し装置。

【 請求項6 】 前記中断時または終了時の情報は、先頭ログの時刻と最終読出しオフセットを含み、前記読出し制御部は、前記先頭ログの時刻に基づいて処理するファイルを前記ファイル検索部によって特定し、前記最終読出しオフセットに基づいて特定したファイル内の読出し位置を復元する、ことを特徴とする請求項5に記載のログ読出し装置。

【 請求項7 】 請求項2乃至6のうちの何れか一項に記載のログ読出し装置を備え、該ログ読出し装置により時系列に沿って読み出されたログデータに基づいて前記ログファイルを作成した装置における稼働分析を行う、ことを特徴とする分析装置。

【 請求項8 】 複数のファイルに格納されたログデータを時系列に沿って読み出すプログラムを記録したコンピュータ読み取り可能な記録媒体において、

前記プログラムは、前記各ファイルの先頭ログデータが持つ時刻情報をそれぞれ取出し、該時刻情報を時系列でソートして前記複数のファイルの読出し順序を決定する、記録媒体。

## 【 発明の詳細な説明】

## 【 0001 】

【 発明の属する技術分野】 本発明は、WWW関連サーバの稼働状況分析で利用するログを読み出すサーバのログ読出し方法、ログ読出し装置、分析装置、及び記録媒体に関するものである。

【 0002 】 近年、インターネット等のネットワークでは、サーバコンピュータに蓄積された種々の情報をクライアントからの要求に応じて提供する検索サービス等のサービスが活用されている。これらのサービスでは、クライアントからWWWサーバへ要求を送信し、その後WWWサーバから応答を受信して結果を表示する形態をとる。このようなサーバにおいて、システムの使用傾向などを分析する稼働状況分析が行われる。この稼働状況分析は、サーバのアクセス記録などを記録したファイル（ログファイル）を利用し、ログを時系列で順番に読み出して行われる。しかし、ログの格納形態（ファイル名、フォーマットなど）はサーバによって異なり、各サーバに対応してログを時系列に読み出すことが難しい。このため、各サーバに容易に対応することができるログ読出し方法が求められている。

## 【 0003 】

【 従来の技術】 従来、コンピュータネットワークによるインターネットにおいては、このインターネット上にWebサーバやProxyサーバ等、複数種類のサーバ（これらはWWW関連サーバであり、これらを総称してインターネットサーバと呼ぶこともある）が接続される。例えば、ユーザはクライアントコンピュータにて直接またはProxyサーバを介して情報提供者のWebサーバ等にアクセスし、このWebサーバに蓄積された種々の情報を閲覧することができる。

【 0004 】 このようなWWW関連サーバの稼働分析では、各サーバが蓄積しているログをデータとして利用することにより、サーバの使用傾向（例えば、Webページのアクセス数や単位時間毎のアクセス数の変化）等の時間的な稼働状況の変化を把握することができる。

## 【 0005 】

【 発明が解決しようとする課題】 ところで、ログファイルの形態は、サーバによって様々であり、単純に1つのファイルから構成されるケースもあれば、規則的な切り替えによって複数のファイルから構成されるケースもある。後者については、ファイルを切り替える度に作成日等の情報を付加したファイル名にて新しいファイルを作成してそのファイルにログデータを記録するタイプのものや、ファイルを切り替える度に使用済みのファイル名をリネームしてログデータを記録するファイルを常に同一

ファイル名にて新しく作成するタイプのものなどがある。このため、複数のファイルで構成されるログファイルからログデータを時系列に読み出すには、サーバ固有のファイル命名規則に沿ってログデータを読み出すファイルの順序を決定し、その順で各ファイルからログデータを読み出していた。

【0006】しかし、サーバによっては、ファイルの命名規則(ファイルのアクセス順序)と格納されたログデータの時系列が一致しないものがある。このように、読み出し方が各サーバの命名規則により決定されるため単純にログデータを時系列で読み出すことができず、ログデータを分析するためには、各サーバそれぞれに対応してログデータの読み出し部を実装する必要がある。

【0007】本発明は上記問題点を解決するためになされたものであって、その目的は格納媒体が複数ファイルで構成されるログについて、サーバ固有の格納媒体生成規則に依存しない、汎用的なログの読み出し方式、ログ読出し装置、分析装置、及び記録媒体を提供することにある。

【0008】

【課題を解決するための手段】上記目的を達成するため、請求項1に記載の発明は、複数のファイルに格納されたログデータを時系列に沿って読み出すログ読出し方法において、前記各ファイルの先頭ログデータが持つ時刻情報をそれぞれ取出し、該時刻情報を時系列でソートして前記複数のファイルの読出し順序を決定する。これにより、ファイルの命名規則に関わらずにログデータを時系列で読み出すことができる。

【0009】請求項2に記載の発明は、複数のファイルに格納されたログデータを時系列に沿って読み出すログ読出し装置において、前記各ファイルのログデータから時刻情報を含む所望のデータを取り出すデータ取出し部と、前記データ取出し部により取り出された各ファイルの先頭ログデータが持つ時刻情報を時系列にソートして次に処理するファイルを検索するファイル検索部と、前記ファイル検索部及びデータ取出し部を、前記ログデータを時系列に沿って読み出すべく制御する読出し制御部と、を備えた。これにより、ファイルの命名規則に関わらずにログデータを時系列で読み出すことができる。

【0010】前記読出し制御部は、請求項3に記載の発明のように、1つのファイルに対する読み出しが終了し、次のファイルを決める段階でその都度前記ファイル検索部を動作させ、前記複数のファイルをソートして読出し順序を決定する。これにより、読出し中にファイルがリネームされても、ログデータを時系列で読み出すことができる。

【0011】前記ファイル検索部は、請求項4に記載の発明のように、前記複数のファイルのうちの2つを順次オープンし、双方のアクセス環境を保持した状態でそれぞれのファイルの先頭ログが持つ時刻情報を比較し、そ

の比較結果に基づいて前記読出し順序を決定する。これにより、比較中にファイルがリネームされてもログデータを時系列で読み出すことができる。

【0012】請求項5に記載の発明のように、前記読出しの中断時またはログデータの終了時の情報を履歴ファイルに保存する履歴記録部を備え、前記読出し制御部は、前記履歴ファイルから読み出した情報に基づいて前記中断時または終了時の読出し位置を復元して読出しを再開する。これにより、読出し再開時における手間を省くことができる。

【0013】前記中断時または終了時の情報は、請求項6に記載の発明のように、先頭ログの時刻と最終読出しオフセットを含み、前記読出し制御部は、前記先頭ログの時刻に基づいて処理するファイルを前記ファイル検索部によって特定し、前記最終読出しオフセットに基づいて特定したファイル内の読出し位置を復元する。これにより、中断から再開までの間にファイルがリネームされてもログデータを時系列で読み出すことができる。

【0014】請求項7に記載の発明によれば、分析装置は、請求項2乃至6のうちの何れか一項に記載のログ読出し装置を備え、該ログ読出し装置により時系列に沿って読み出されたログデータに基づいて前記ログファイルを作成した装置における稼働分析を行う。これにより、ファイルの命名規則に関わらずにログデータを時系列で読み出して分析を行うことができる。

【0015】請求項8に記載の発明は、複数のファイルに格納されたログデータを時系列に沿って読み出すプログラムを記録したコンピュータ読み取り可能な記録媒体において、前記プログラムは、前記各ファイルの先頭ログデータが持つ時刻情報をそれぞれ取出し、該時刻情報を時系列でソートして前記複数のファイルの読出し順序を決定する。これにより、ファイルの命名規則に関わらずにログデータを時系列で読み出すことができる。

【0016】尚、上記各請求項において、複数のファイルが存在するパスがワイルドカードを用いて指定される。これにより、複数のファイルの保存位置に関係なくログデータを読み出すことができる。

【0017】また、取り出すデータの種を示すキーワードが記録形式として指定され、該記録形式に基づいて前記各ファイルのログデータから所望のデータを取り出すようにしてもよい。これにより、固定されたログファイルの記録形式に依存しない読出しが可能となる。

【0018】

【発明の実施の形態】(第一実施形態)以下、本発明を具体化した第一実施形態を図1～図4に従って説明する。

【0019】図1は、システムの構成及び接続形態を概略的に示すブロック図である。このシステムは、クライアントコンピュータ(以下、単にクライアントという)11a, 11b, Proxyサーバ12, Webサーバ

13、分析用コンピュータ14を含む。クライアント11aはネットワーク16に直接的に接続され、クライアント11bはProxyサーバ12を介してネットワークに接続される。Webサーバ13及び分析用コンピュータ14はネットワーク16に直接的に接続される。

【0020】Webサーバ13は、ユーザが操作するクライアント11a、11bからの要求にตอบสนองしてWebページ等のコンテンツを提供するサービス機能を持ち、そのサービス提供の状況をログファイル17に格納する。尚、ログファイル17はWebサーバ13がログを格納する格納媒体を機能的に示したものであり、Webサーバ13が持つ命名規則に従って作成される複数のファイルにより構成されるものである。

【0021】分析用コンピュータ14は、ネットワーク16を介してWebサーバ13のログファイル17からログデータを時系列に読み出し、Webサーバ13の稼働状況を分析する機能を持つ。

【0022】図2は、分析用コンピュータ14に対応する情報処理装置(コンピュータ)のハードウェア構成を概略的に示すブロック図である。分析用コンピュータ14が提供する機能は、このような汎用的な目的で使用される計算機上で実行するコンピュータプログラムにより実現する。尚、クライアント11a、11b、Proxyサーバ12、Webサーバ13も同様に汎用的な計算機上で実行するコンピュータプログラムにより実現する。

【0023】情報処理装置は、中央処理装置(CPU)21を含み、このCPU21は、メモリ22と、入力装置23と、出力装置24と、記憶装置25と、ドライブ装置26と、通信装置27とをそれに接続された状態で有している。

【0024】CPU21は、メモリ22を利用してプログラムを実行し、ログデータの読み出し、稼働分析等の機能を実現する。メモリ22には、機能の提供に必要なプログラムとデータが格納され、メモリ22としては、通常、キャッシュ・メモリ、システム・メモリおよびディスプレイ・メモリを含む。

【0025】入力装置23は、キーボードおよびマウス装置(図示せず)を含み、ユーザからの要求や指示の入力に用いられる。また、出力装置24は、表示装置(VDT)またはモニタおよびプリンタを含み、サービス画面、分析結果等の出力に用いられる。

【0026】記憶装置25は、通常、磁気ディスク装置、光ディスク装置、光磁気ディスク装置を含む。この記憶装置25には、上述のプログラムとデータが格納され、CPU21は、入力装置23による指示にตอบสนองしてプログラム、データをメモリ22へ転送し、それを実行する。

【0027】ドライブ装置26は、記録媒体28を駆動し、その記憶内容にアクセスする。記録媒体28として

は、メモリカード、フロッピーディスク、光ディスク(CD-ROM、DVD-ROM、...)、光磁気ディスク(MO、MD、...)等、任意のコンピュータ読み取り可能な記録媒体を使用することができる。この記録媒体28に、上述のプログラム、データを格納しておき、必要に応じて、メモリ22にロードして使用することもできる。

【0028】通信装置27は、ネットワーク16への接続に用いられ、通信に伴うデータ変換等を行って、他の情報処理装置(クライアント11a、11b、Proxyサーバ12、Webサーバ13)との間でプログラムやデータの送受信を行う。また、情報処理装置は、外部の情報提供者のデータベース等からネットワークを介して、上述のプログラムとデータを受け取り、必要に応じて、メモリ22にロードして使用することもできる。

【0029】図3は、分析用コンピュータ14とログファイル17を機能的に示すブロック図である。ログファイル17は、複数(図3では4つ)のファイル17a、17b、17c、17dから構成されている。各ファイル17a~17dは、それぞれファイル名がWebサーバ13の命名規則により「ファイル1」、「ファイル2」、「ファイル3」、「ファイル4」と命名されている。

【0030】尚、本実施形態では、ログファイル17は、以下の条件により作成されている。

ーファイルはテキストファイルである。

ーファイルは同一ディレクトリ配下に存在する。

ーログはテキスト行一行に一回分が記録される。

【0031】分析用コンピュータ14は、ログ分析部31とデータ読出し部32を含み、そのデータ読出し部32は、読出し制御部33、ファイル検索部34、データ取出し部35を備えている。

【0032】ログ分析部31はその分析処理時にデータ読出し部32を介してログファイル17から時系列的に読み出したログデータに基づいてWebサーバ13の稼働状況を分析する。ログ分析部31は、分析しようとするログの時刻(分析開始時刻)等の読出し情報を指定してデータ読出し部32を呼び出す。

【0033】データ読出し部32は、読出し情報とWebサーバ13に対応して設定されるログ情報に基づいて各ファイル17a~17dの読出し順序を決定し、その決定した読出し順序に従って読み出したログデータをログ分析部31に出力する。

【0034】ログ情報は、Webサーバの制約事項(Webサーバ毎に決められている事項)やそれを運用するユーザが決定する事項等、である。尚、本実施形態ではログ情報として以下の事項を外部から与える。

ーファイル群が存在するディレクトリのパス。

ーファイル名パターン(可変部分をワイルドカードで指定)。

ーファイル内のログの記録形式(データ種類を示すキー

ワードの並び)。

【0035】読出し制御部33は、時系列読出しを制御する役割を担い、処理の経過に従って適宜ファイル検索部34又はデータ取出し部35を呼び出す。ファイル検索部34は、外部から与えられるログ情報(「ディレクトリのパス」と「ファイル名パターン」)に基づいて、次の処理するファイルを探索し、その探索したファイルをオープンした環境を読出し制御部33に引き渡す役割を担う。

【0036】データ取出し部35は、読出し制御部33及びファイル検索部34から呼び出され、外部から与えられるログ情報(「記録形式」)に基づいて、ファイルから読み出した1行分のログデータから時刻情報等の各種分析用データを取り出し、それらを呼び出し元に出力する。

【0037】詳述すると、読出し制御部33は、分析部31から指定される分析開始時刻を基点時刻として指定してファイル検索部34を呼び出す。呼び出されたファイル検索部34は、「ディレクトリのパス」と「ファイル名パターン」に基づいてファイルを順次オープンし、オープン環境にあるファイルからログデータを1行分読出し、データ取出し部35を呼び出す。呼び出されたデータ取出し部35は、ログデータ1行分から「記録形式」に基づいて取り出した時刻情報等の各種分析用データを読み出し元であるファイル検索部34に出力する。

【0038】図4は、ログの「記録形式」を示す説明図である。この例では、ログデータの一行には、アクセス時刻、アクセス元情報(クライアント情報)、被アクセス情報(アクセスされたファイル名)、データ量、その処理時間が格納されており、これらの情報に対応するキーワードの並びが記録形式41として外部から指定される。これにより、サーバ固有のログの記録形式に依存しない読出しを可能としている。この記録形式41は、Webサーバ13に対応して一度すれば良い。

【0039】ファイル検索部34は、各ファイルの先頭ログデータが持つ時刻情報に基づいて時系列にファイルをソートして読出し順序を決定する。そして、ファイル検索部34は、決定した読出し順序に基づいて次に処理するファイルをオープンし、そのファイルのオープン環境を読出し制御部33に引き渡す。

【0040】具体的には、ファイル検索部34は2つのファイルをオープンし、双方のアクセス環境(オープン環境)を保持した状態でそれらの先頭ログデータが持つ時刻情報を比較する。そして、ファイル検索部34は、基点時刻以降で基点時刻に最も近い時刻情報を持つログデータを格納した一方のファイルのオープン環境を保持し、他方のファイルをクローズして次のファイルをオープンする。このように、ファイル検索部34は、全てのファイルの先頭ログデータが持つ時刻情報を比較し、基点時刻以降で基点時刻に最も近い時刻情報を持つログデ

ータを格納したファイルのオープン環境を読出し制御部33に受け渡す。

【0041】尚、ファイル検索部34は、1つのファイルに対する読出しが終了し、次に処理するファイルを決定する段階でソート処理をその都度実行する。これにより、読出し中に実施されたファイルのリネーム処理に対応する。また、ファイル検索部34は、ソート時に2つのファイルの比較を双方のアクセス環境を保持した状態(双方のファイルをオープンした状態)で実行することで、比較中のリネーム処理に対応する。

【0042】即ち、Webサーバ13は、アクセスがある毎にログファイル17を更新する。従って、ソート処理を1度だけ実行してファイルの読出し順序を決定する方法では、リネームされてファイル名が変更されると読出し順序が変わってしまい、時系列にログデータを読み出すことができなくなる。また、双方のファイルをクローズして比較する方法では、比較中にリネーム処理が行われると、比較結果に基づいてオープンするファイル名が変わる又はそのファイルが無くなってしまふことがあり、やはり時系列にログデータを確実に読み出すことができなくなるからである。

【0043】読出し制御部33は、オープン環境を受け取り、オープン環境にあるファイルからログデータを1行分読出し、データ取出し部35を呼び出す。呼び出されたデータ取出し部35は、ログデータ1行分から「記録形式」に基づいて取り出した時刻情報等の各種分析用データを読み出し元である読出し制御部33に出力する。読出し制御部33は、引き継いだオープン環境がファイルの終端に到達するまで(ファイルのログデータが無くなるまで)データ取出し部35を呼び出してそのファイルの全てのログデータから取り出した各種分析用データを受け取り、それらをログ分析部31に出力する。

【0044】このように、1つのファイルのログデータを全てについて処理を終了すると、読出し制御部33は、次のファイルの読出しを行うべく例えば先のファイルの最終ログデータの時刻情報を基点時刻として再度ファイル検索部34を呼び出す。

【0045】次に、ファイル検索について詳述する。尚、ここでは、UNIXにおけるC言語による実装として示す。例えば、ファイル2、ファイル3、ファイル4、ファイル1の順に時系列になっており、システムコール `opendir()` 関数でディレクトリをオープンした場合に、`readdir()` 関数でディレクトリエントリがファイル1、ファイル2、ファイル3、ファイル4の順で読み出される。この時、ファイル検索部34が実施する最初に処理するファイルを探索する処理は、以下のとおりである。

【0046】(1) `fopen()` 関数で「ファイル1」をオープンし、`fgets()` 関数で先頭ログのテキスト行を得、データ取出し部35を呼び出して先頭ログの時刻情報を

取り出す。

【0047】(2) fopen()関数で「ファイル2」をオープンし、fgets()関数で先頭ログのテキスト行を得、データ取出し部35を呼び出して先頭ログの時刻情報を取り出す。

【0048】(3) 時刻情報を比較すると「ファイル2」の方が古いので、「ファイル2」のオープン環境をそのままにして、fclose()関数で「ファイル1」のオープン環境をクローズする。

【0049】(4) fopen()関数で「ファイル3」をオープンし、fgets()関数で先頭ログのテキスト行を得、データ取出し部35を呼び出して先頭ログの時刻情報を取り出す。

【0050】(5) 時刻情報を比較すると「ファイル2」の方が古いので、「ファイル2」のオープン環境をそのままにして、fclose()関数で「ファイル3」のオープン環境をクローズする。

【0051】(6) fopen()関数で「ファイル4」をオープンし、fgets()関数で先頭ログのテキスト行を得、データ取出し部35を呼び出して先頭ログの時刻情報を取り出す。

【0052】(7) 時刻情報を比較すると「ファイル2」の方が古いので、「ファイル2」のオープン環境をそのままにして、fclose()関数で「ファイル4」のオープン環境をクローズする。

【0053】(8) 全てのファイルについて検索処理が終了したので、該当ファイル名を「ファイル2」とし、「ファイル2」のオープン環境をそのまま読み出し制御部33へ引き渡す。

【0054】以上記述したように、本実施の形態によれば、以下の効果を奏する。

(1) データ読み出し部32は、各ファイル17a~17dのログデータから時刻情報を含む所望のデータを取り出すデータ取出し部35と、データ取出し部35により取り出された各ファイル17a~17dの先頭ログデータが持つ時刻情報を時系列にソートして次に処理するファイルを検索するファイル検索部34と、ファイル検索部34及びデータ取出し部35を、ログデータを時系列に沿って読み出すべく制御する読み出し制御部33と、を備えた。これにより、ファイルの命名規則に関わらずにログデータを時系列で読み出すことができる。

【0055】(2) 読み出し制御部33は、1つのファイルに対する読み出しが終了し、次のファイルを決める段階でその都度ファイル検索部34を動作させ、複数のファイルをソートして読み出し順序を決定する。これにより、読み出し中にファイルがリネームされても、ログデータを時系列で読み出すことができる。

【0056】(3) ファイル検索部34は、複数のファイルのうちの2つを順次オープンし、双方のアクセス環境を保持した状態でそれぞれのファイルの先頭ログが持

つ時刻情報を比較し、その比較結果に基づいて読み出し順序を決定する。これにより、比較中にファイルがリネームされてもログデータを時系列で読み出すことができる。

【0057】(4) ファイル検索部34は、複数のファイルが存在するパスがワイルドカードを用いて指定されたパスに存在する複数のファイル17a~17dをソートする。これにより、複数のファイルの保存位置に関係なくログデータを読み出すことができる。

【0058】(5) データ取出し部35は、取り出すデータの種別を示すキーワードが記録形式として指定され、該記録形式に基づいて各ファイルのログデータから所望のデータを取り出す。これにより、固定されたログファイルの記録形式に依存しない読み出しを行うことができる。

【0059】(第二実施形態)以下、本発明を具体化した第二実施形態を図5に従って説明する。尚、説明の便宜上、第一実施形態と同様の構成については同一の符号を付してその説明を一部省略する。

【0060】図5は、分析用コンピュータ14aとログファイル17を機能的に示すブロック図である。尚、分析用コンピュータ14aに対応する情報処理装置(コンピュータ)のハードウェア構成は、第一実施形態の分析用コンピュータ14と同じである(図2参照)ため、図面及び詳細な説明を省略する。

【0061】分析用コンピュータ14aは、ログ分析部31aとデータ読み出し部32aを含み、そのデータ読み出し部32aは、読み出し制御部33a、ファイル検索部34a、データ取出し部35、履歴記録部51を備えている。

【0062】ログ分析部31aはその分析処理時にデータ読み出し部32aを介してログファイル17から時系列的に読み出したログデータに基づいてWebサーバ13の稼働状況を分析する。

【0063】ログ分析部31aは、分析しようとするログの時刻(分析開始時刻)等の読み出し情報又は読み出しの再開又は中断を指定してデータ読み出し部32aを呼び出す。データ読み出し部32aは、読み出し情報とWebサーバ13に対応して設定されるログ情報に基づいて各ファイル17a~17dの読み出し順序を決定し、その決定した読み出し順序に従って読み出したログデータをログ分析部31aに出力する。また、データ読み出し部32aは、読み出しの再開又は中断指示に基づいて、先に読み出し処理した最後のログデータの次のログデータからの読み出し、または読み出しの中断処理を行う。

【0064】読み出し制御部33aは、読み出しの中断指示を受け取るか、またはログファイル17の全てのログデータを読み出すと、履歴記録部51を制御してその時点の状態を保存する。具体的には、読み出し制御部33aは、その時の「先頭ログの時刻」及びftell()関数によ

って得られる「最終読出しオフセット」を履歴記録部5 1 に出力し、履歴記録部5 1 はそれらを履歴ファイル5 2 に記録する。尚、履歴を元にして読出しを再開できるように、ファイル検索部3 4 a は指定した時刻に一致する先頭ログを持つファイルを検索する機能を備える。

【0065】読出し制御部3 3 a は、読出しの再開指示を受け取ると、まず、履歴記録部5 1 を介して履歴ファイル5 2 から「先頭ログの時刻」及び「最終読出しオフセット」を読出し、その「先頭ログの時刻」を基点時刻として指定してファイル検索部3 4 a を呼び出す。ファイル検索部3 4 a は、指定された基点時刻に対応するファイルのオープン環境を読出し制御部3 3 a に引き渡す。

【0066】読出し制御部3 3 a は、次に、「最終読出しオフセット」でファイル内の読出し位置を復元し、その復元位置からデータ取出し部3 5 を制御して読出しを再開する。これにより、読出しの中断指示又は所定の集計時間に満たないログデータの集計を再開する機能を提供する。

【0067】即ち、ログデータの読出し中に分析部3 1 a からの指示によって読出しが中断する場合がある。また、読み出したログデータが所定の集計時間（例えば1 時間、1 日など）に満たない（所定の集計時間内のログデータの収集が済んでいない）場合がある。これらの場合、読出しを再開するファイルを検索するのに時間がかかったり、「先頭ログの時刻」だけでファイル検索を行うと、先に読み出したログデータを再度読み込んで集計したり、ログデータを2 度集計してしまう等することがある。このため、履歴ファイル5 2 に「先頭ログの時刻」及び「最終読出しオフセット」を記録し、これらに基づいて読出しを再開することで、ファイル検索の手間や、集計のだぶりを防ぐ。また、「先頭ログの時刻」で読出しを再開するファイルを特定することで、中断から再開までの間における該当ファイルのリネームに対応する。即ち、ファイル名が変更されても、「先頭ログの時刻」が変えられない以上、同じ内容を持つファイルを容易にオープンすることができる。

【0068】以上記述したように、本実施の形態によれば、第一実施形態の効果に加えて以下の効果を奏する。

（1）データ読出し部3 2 a は、読出しの中断時またはログデータの終了時の情報を履歴ファイル5 2 に保存する履歴記録部5 1 を備え、読出し制御部3 3 a は、履歴ファイル5 2 から読み出した情報に基づいて中断時または終了時の読出し位置を復元して読出しを再開する。これにより、読出し再開時における手間を省くことができ、ログの時分割的な読出しや、EOF (end of file) からの追加分の読出しを容易に行うことができる。

【0069】（2）履歴ファイル5 2 に格納される中断時または終了時の情報は、「先頭ログの時刻」と「最終読出しオフセット」を含み、読出し制御部3 3 a は、

「先頭ログの時刻」に基づいて処理するファイルをファイル検索部3 4 a によって特定し、「最終読出しオフセット」に基づいて特定したファイル内の読出し位置を復元する。これにより、中断から再開までの間にファイルがリネームされてもログデータを時系列で読み出すことができる。

【0070】尚、前記各実施形態は、以下の態様に変更してもよい。

○上記各実施形態では、Web サーバ1 3 と分析用コンピュータ1 4 を別々の構成としたが、同一の計算機にてWeb ページ等のサービス機能とログの読出し及び分析機能を提供する構成としても良い。

【0071】○上記各実施形態では、ログデータの読出し部3 2, 3 2 a を分析用コンピュータ1 4, 1 4 a に実装したが、これをWeb サーバ1 3 に実装し、時系列的に読出したログデータを順次、或いは所定のまとまりを持って分析用コンピュータ1 4, 1 4 a に送信する構成としても良い。

【0072】○上記各実施形態では、Web サーバ1 3 のログを読み出す読出し部3 2, 3 2 a に具体化したのが、図1 のProxy サーバ1 2 や、他のWWW関連サーバのログデータを読み出す読出し部3 2, 3 2 a に具体化して実施しても良い。WWW関連サーバとしては、Web サーバのようにネットワークに接続されたクライアントにデータベース、メール等のサービスを提供するもの、Proxy サーバのようにクライアントをネットワークに接続するためのサービスを提供するものを含む。

【0073】○上記各実施形態では、C 言語により実装した機能（プログラム）について説明したが、他の言語を用いて作成したプログラムにより機能を実装するようにしてもよい。

【0074】以上の実施形態をまとめ、本発明の構成に関する以下の事項を開示する。

（1）前記ソートをし、1 つのファイルに対する読み出しが終了し、次のファイルを決める段階でその都度実行する、ことを特徴とする請求項1 に記載のログ読出し方法。これにより、読出し中にファイルがリネームされても、ログデータを時系列で読み出すことができる。

【0075】（2）前記ソート時に前記複数のファイルのうちの2 つを順次オープンし、双方のアクセス環境を保持した状態でそれぞれのファイルの先頭ログデータが持つ時刻情報を比較し、その比較結果に基づいて前記読出し順序を決定する、ことを特徴とする請求項1 に記載のログ読出し方法。これにより、比較中にファイルがリネームされてもログデータを時系列で読み出すことができる。

【0076】（3）前記読出しの中断時またはログデータの終了時の情報を保存し、該情報を読み出して前記中断時または終了時の読出し位置を復元して読出しを再



13

開する、ことを特徴とする請求項1に記載のログ読出し方法。これにより、読出し再開時における手間を省くことができる。

【0077】(4) 前記情報に含まれる先頭ログの時刻に基づいて処理するファイル特定し、前記情報に含まれる最終読み出しオフセットに基づいて前記特定したファイル内の読み出し位置を復元する、ことを特徴とする上記(3)に記載のログ読出し方法。これにより、中断から再開までの間にファイルがリネームされてもログデータを時系列で読み出すことができる。

【0078】(5) 前記プログラムは、前記ソートを、1つのファイルに対する読み出しが終了し、次のファイルを決める段階でその都度実行する、請求項8に記載の記録媒体。

【0079】(6) 前記プログラムは、前記複数のファイルのうちの2つを順次オープンし、双方のアクセス環境を保持した状態でそれぞれのファイルの先頭ログデータが持つ時刻情報を比較し、その比較結果に基づいて前記読出し順序を決定する、請求項8に記載の記録媒体。

【0080】(7) 前記プログラムは、前記読出しの中断時またはログデータの終了時の情報を保存し、該情報を読み出して前記中断時または終了時の読出し位置を復元して読出しを再開するステップを実行する、請求項8に記載の記録媒体。

【0081】(8) 前記プログラムは、前記情報に含まれる先頭ログの時刻に基づいて処理するファイル特定し、前記情報に含まれる最終読み出しオフセットに基

14

づいて前記特定したファイル内の読み出し位置を復元するステップを実行する、上記(7)に記載の記録媒体。

【0082】

【発明の効果】以上詳述したように、本発明によれば、複数のファイルの先頭ログデータが持つ時刻情報をソートしてファイルの読出し順序を決定するようにしたこと、WWW関連サーバ(Webサーバ、Proxyサーバなど)の稼働状況分析において、サーバの固有情報に関わらずにログデータを時系列で読出しを行うことができる。

【図面の簡単な説明】

【図1】 システム構成を概略的に示すブロック図である。

【図2】 分析用コンピュータのハードウェア構成を概略的に示すブロック図である。

【図3】 第一実施形態の分析用コンピュータの機能ブロック回路図である。

【図4】 ログの記録形式を示す説明図である。

【図5】 第二実施形態の分析用コンピュータの機能ブロック回路図である。

【符号の説明】

17, 17a~17d ログファイル

33, 33a 読出し制御部

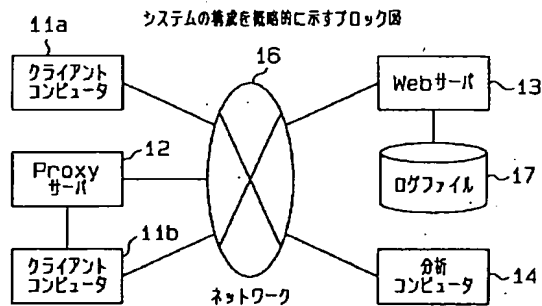
34, 34a ファイル検索部

35 データ取出し部

51 履歴記録部

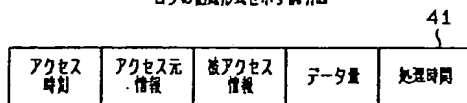
52 履歴ファイル

【図1】

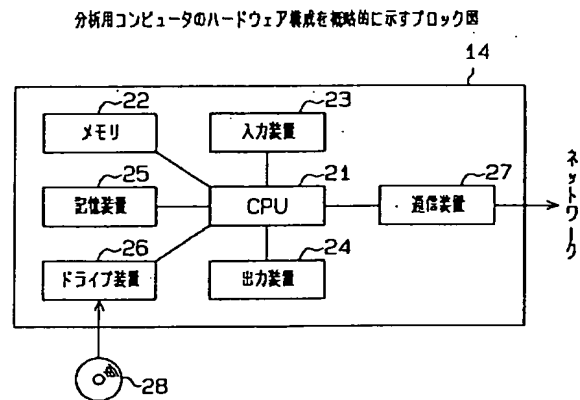


【図4】

ログの記録形式を示す説明図

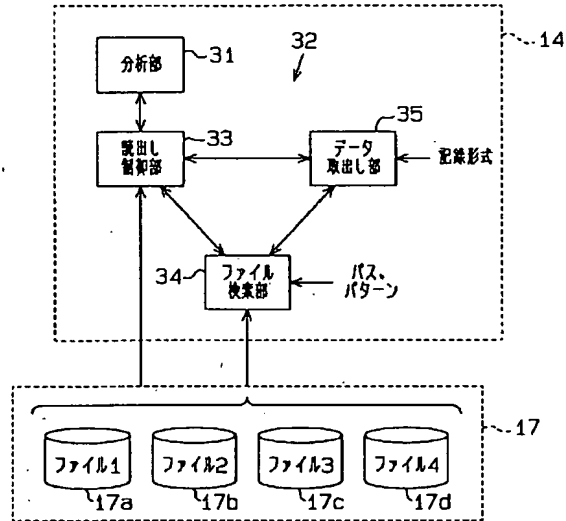


【図2】



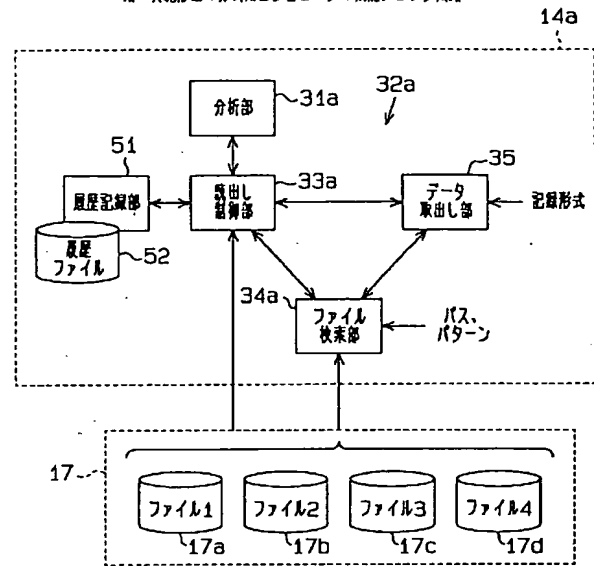
【 図3 】

第一実施形態の分析用コンピュータの機能ブロック回路図



【 図5 】

第二実施形態の分析用コンピュータの機能ブロック回路図



フロント ページの続き

(72)発明者 山崎 善広  
愛知県名古屋市東区葵1 丁目16番38号 株  
式会社富士通愛知エンジニアリング内

F ターム (参考) 5B042 GA12 GB02 GC08 HH20 MA12  
MA14 MC40  
5B075 ND20 NR15 PP02 PP03  
5B082 EA07 GA03

BEST AVAILABLE COPY